

# Plan de análisis de capacidad

MISW 4204 – Desarrollo de software en la nube

Simón Buriticá

Jhonn Sebastian Calderón Bravo

Diego Andrés Naranjo Ríos

Juan Pablo Rodríguez García

Universidad de los Andes - 2024

## Contexto:

La funcionalidad de cargar y editar videos es la de mayor valor para los pilotos que deseen participar en la competencia, al igual que la de mayor complejidad debido a la cantidad de subprocesos y recursos que consume. Por tal razón, poder determinar la capacidad de procesamiento de tareas de edición en simultaneo y las condiciones de operación ideales que permitan obtener la máxima salida de videos (throughput) es crucial para ofrecer a los usuarios la mejor disponibilidad y latencia posible.

## Objetivos:

1. Determinar la cantidad óptima de videos a procesar en simultaneo
2. Calcular el tiempo promedio de procesamiento de un video desde su carga en minutos
3. Medir el porcentaje de error de solicitudes para varios escenarios de procesamiento de video

## Criterios de aceptación:

Para el contexto de las pruebas, se definen los siguientes valores como los esperados para la configuración óptima:

- Disponibilidad del  $98\% \pm 2\%$
- Uso de CPU de aproximadamente 90%
- Error en las peticiones menor a 1%

## Entorno de prueba:

Entorno	Máquina virtual con 2 vCPU + 2GB Ram + 20 GB Almacenamiento
Sistema operativo	Sistema operativo Ubuntu 22
Base de datos	PostgreSQL en contenedor
Lenguaje API / Framework	Python / Flask
Método de despliegue	Contenedores de Docker
Método de comunicación	HTTP remoto. 80-100ms de latencia esperada
Software de pruebas	JMeter

Tabla 1 - Condiciones del entorno de pruebas

Artefactos de las pruebas
Pool de 4 videos de entre 20mb a 25mb de peso y 4 a 5 minutos de duración
Lista de 100 usuarios registrados en la plataforma
Tablas de datos de videos y tareas vacías

Tabla 2 - Artefactos de prueba

## Metodología

Paso	Descripción
1	Se preselecciona un usuario aleatorio del pool de usuarios
2	Autenticar con los datos del usuario para obtener un token de autorización
3	Realizar la petición para cargar un video aleatorio con el token conseguido
4	Validar cada medio segundo el estado de la tarea hasta que sea finalizado
5	Medir el tiempo de ejecución de la prueba

Tabla 3 - Escenario de prueba por cada instancia de JMeter

En la tabla 3 se presenta el escenario de pruebas que se ejecutará por cada instancia de JMeter. A partir de este punto se escalará el número de usuarios concurrentes que deciden hacer la solicitud a la plataforma. **El escalamiento de las pruebas se realizará entre 1 a 10 usuarios concurrentes, replicando cada la prueba 3 veces por cada conjunto de usuarios, para un total de 30 pruebas.** En caso de que no se encuentre una variación significativa, o que la máquina todavía recursos disponibles, **se aumentará el número de instancias concurrentes hasta 20 usuarios.**

## Análisis de escenarios

A partir de la metodología descrita, se piensa dar respuesta a los siguientes escenarios:

1. Comportamiento de la plataforma ante la carga en simultaneo de videos
2. Capacidad para la edición de videos en batches

Estos escenarios se pueden determinar a partir del escenario de prueba descrito usando el porcentaje de fallas y el tiempo medio de edición de videos respectivamente. Se espera que para ambos escenarios se genere una gráfica que permita relacionar el throughput del sistema y el tamaño del batch de videos, al igual que encontrar cuál es la capacidad de la capa web para atender a los usuarios